PCT10

RAW SEQUENCE LISTING   DATE: 02/08/2002
PATENT APPLICATION: US/10/030,298   TIME: 10:52:14

Input Set : A:\SO043PCTSEQ.txt
Output Set: N:\CRF3\02082002\J030298.raw   **ENTERED**

```
   3 <110> APPLICANT: Shionogi & Co., Ltd
   5 <120> TITLE OF INVENTION: Drug for inhibiting myometrial contraction
   7 <130> FILE REFERENCE: SO043PCT
C-->  9 <140> CURRENT APPLICATION NUMBER: US/10/030,298
C--> 10 <141> CURRENT FILING DATE: 2001-12-21
  11 <150> PRIOR APPLICATION NUMBER: JP P1999-177548
  12 <151> PRIOR FILING DATE: 1999-06-23
  13 <150> PRIOR APPLICATION NUMBER: JP P2000-79171
  14 <151> PRIOR FILING DATE: 1999-03-21
  16 <160> NUMBER OF SEQ ID NOS: 6
  18 <170> SOFTWARE: PatentIn Ver. 2.0
  20 <210> SEQ ID NO: 1
  21 <211> LENGTH: 1457
  22 <212> TYPE: DNA
  23 <213> ORGANISM: Homo sapiens
  25 <220> FEATURE:
  26 <221> NAME/KEY: CDS
  27 <222> LOCATION: (165)..(719)
  29 <220> FEATURE:
W--> 30 <221> NAME/KEY: mat peptide
  31 <222> LOCATION: (447)..(602)
  33 <400> SEQUENCE: 1
  34 ggcacgagct ggatagaaca gctcaagcct tgccacttcg ggcttctcac tgcagctggg 60
  36 cttggacttc ggagttttgc cattgccagt gggacgtctg agactttctc cttcaagtac 120
  38 ttggcagatc actctcttag cagggtctgc gcttcgcagc cggg atg aag ctg gtt 176
  39                                                  Met Lys Leu Val
  41 tcc gtc gcc ctg atg tac ctg ggt tcg ctc gcc ttc cta ggc gct gac 224
  42 Ser Val Ala Leu Met Tyr Leu Gly Ser Leu Ala Phe Leu Gly Ala Asp
  43 -90                 -85                 -80                 -75
  45 acc gct cgg ttg gat gtc gcg tcg gag ttt cga aag aag tgg aat aag 272
  46 Thr Ala Arg Leu Asp Val Ala Ser Glu Phe Arg Lys Lys Trp Asn Lys
  47                 -70                 -65                 -60
  49 tgg gct ctg agt cgt ggg aag agg gaa ctg cgg atg tcc agc agc tac 320
  50 Trp Ala Leu Ser Arg Gly Lys Arg Glu Leu Arg Met Ser Ser Ser Tyr
  51             -55                 -50                 -45
  53 ccc acc ggg ctc gct gac gtg aag gcc ggg cct gcc cag acc ctt att 368
  54 Pro Thr Gly Leu Ala Asp Val Lys Ala Gly Pro Ala Gln Thr Leu Ile
  55         -40                 -35                 -30
  57 cgg ccc cag gac atg aag ggt gcc tct cga agc ccc gaa gac agc agt 416
  58 Arg Pro Gln Asp Met Lys Gly Ala Ser Arg Ser Pro Glu Asp Ser Ser
  59     -25                 -20                 -15
  61 ccg gat gcc gcc cgc atc cga gtc aag cgc tac cgc cag agc atg aac 464
  62 Pro Asp Ala Ala Arg Ile Arg Val Lys Arg Tyr Arg Gln Ser Met Asn
```

RAW SEQUENCE LISTING   DATE: 02/08/2002
PATENT APPLICATION: US/10/030,298   TIME: 10:52:14

Input Set : A:\SO043PCTSEQ.txt
Output Set: N:\CRF3\02082002\J030298.raw

```
 63 -10                   -5              -1  1               5
 65 aac ttc cag ggc ctc cgg agc ttt ggc tgc cgc ttc ggg acg tgc acg   512
 66 Asn Phe Gln Gly Leu Arg Ser Phe Gly Cys Arg Phe Gly Thr Cys Thr
 67           10              15              20
 69 gtg cag aag ctg gca cac cag atc tac cag ttc aca gat aag gac aag   560
 70 Val Gln Lys Leu Ala His Gln Ile Tyr Gln Phe Thr Asp Lys Asp Lys
 71           25              30              35
 73 gac aac gtc gcc ccc agg agc aag atc agc ccc cag ggc tac ggc cgc   608
 74 Asp Asn Val Ala Pro Arg Ser Lys Ile Ser Pro Gln Gly Tyr Gly Arg
 75        40              45              50
 77 cgg cgc cgg cgc tcc ctg ccc gag gcc ggc ccg ggt cgg act ctg gtg   656
 78 Arg Arg Arg Arg Ser Leu Pro Glu Ala Gly Pro Gly Arg Thr Leu Val
 79 55              60              65                  70
 81 tct tct aag cca caa gca cac ggg gct cca gcc ccc ccg agt gga agt   704
 82 Ser Ser Lys Pro Gln Ala His Gly Ala Pro Ala Pro Pro Ser Gly Ser
 83                 75              80              85
 85 gct ccc cac ttt ctt taggatttag gcgcccatgg tacaaggaat agtcgcgcaa   759
 86 Ala Pro His Phe Leu
 87                 90
 89 gcatcccgct ggtgcctccc gggacgaagg acttcccgag cggtgtgggg accgggctct 819
 91 gacagccctg cggagaccct gagtccggga ggcaccgtcc ggcggcgagc tctggctttg 879
 93 caagggcccc tccttctggg ggcttcgctt ccttagcctt gctcaggtgc aagtgcccca 939
 95 gggggcgggg tgcagaagaa tccgagtgtt tgccaggctt aaggagagga gaaactgaga 999
 97 aatgaatgct gagacccccg gagcaggggt ctgagccaca gccgtgctcg cccacaaact 1059
 99 gatttctcac ggcgtgtcac cccaccaggg cgcaagcctc actattactt gaactttcca 1119
101 aaacctaaag aggaaaagtg caatgcgtgt tgtacataca gaggtaacta tcaatatTta 1179
103 agtttgttgc tgtcaagatt ttttttgtaa cttcaaatat agagatattt ttgtacgtta 1239
105 tatattgtat taagggcatt ttaaaagcaa ttatattgtc ctcccctatt ttaagacgtg 1299
107 aatgtctcag cgaggtgtaa agttgttcgc cgcgtggaat gtgagtgtgt ttgtgtgcat 1359
109 gaaagagaaa gactgattac ctcctgtgtg gaagaaggaa acaccgagtc tctgtataat 1419
111 ctatttacat aaaatgggtg atatgcgaac agcaaacc                         1457
114 <210> SEQ ID NO: 2
115 <211> LENGTH: 185
116 <212> TYPE: PRT
117 <213> ORGANISM: Homo sapiens
119 <400> SEQUENCE: 2
120 Met Lys Leu Val Ser Val Ala Leu Met Tyr Leu Gly Ser Leu Ala Phe
121                 -90             -85             -80
123 Leu Gly Ala Asp Thr Ala Arg Leu Asp Val Ala Ser Glu Phe Arg Lys
124             -75             -70             -65
126 Lys Trp Asn Lys Trp Ala Leu Ser Arg Gly Lys Arg Glu Leu Arg Met
127          -60             -55             -50
129 Ser Ser Ser Tyr Pro Thr Gly Leu Ala Asp Val Lys Ala Gly Pro Ala
130      -45             -40             -35
132 Gln Thr Leu Ile Arg Pro Gln Asp Met Lys Gly Ala Ser Arg Ser Pro
133 -30             -25             -20                 -15
135 Glu Asp Ser Ser Pro Asp Ala Ala Arg Ile Arg Val Lys Arg Tyr Arg
136             -10             -5              -1  1
138 Gln Ser Met Asn Asn Phe Gln Gly Leu Arg Ser Phe Gly Cys Arg Phe
```

RAW SEQUENCE LISTING   DATE: 02/08/2002
PATENT APPLICATION: US/10/030,298   TIME: 10:52:14

Input Set : A:\SO043PCTSEQ.txt
Output Set: N:\CRF3\02082002\J030298.raw

```
139            5                    10                   15
141 Gly Thr Cys Thr Val Gln Lys Leu Ala His Gln Ile Tyr Gln Phe Thr
142        20                   25                   30
144 Asp Lys Asp Lys Asp Asn Val Ala Pro Arg Ser Lys Ile Ser Pro Gln
145  35                   40                   45                   50
147 Gly Tyr Gly Arg Arg Arg Arg Arg Ser Leu Pro Glu Ala Gly Pro Gly
148                 55                   60                   65
150 Arg Thr Leu Val Ser Ser Lys Pro Gln Ala His Gly Ala Pro Ala Pro
151             70                   75                   80
153 Pro Ser Gly Ser Ala Pro His Phe Leu
154            85                   90
157 <210> SEQ ID NO: 3
158 <211> LENGTH: 1493
159 <212> TYPE: DNA
160 <213> ORGANISM: Sus scrofa
162 <220> FEATURE:
163 <221> NAME/KEY: CDS
164 <222> LOCATION: (148)..(711)
166 <220> FEATURE:
```
W--> 167 **<221> NAME/KEY: mat peptide**
```
168 <222> LOCATION: (430)..(585)
170 <400> SEQUENCE: 3
171 gcggaacagc tcgagccttg ccacctctag tttcttacca cagcttggac gtcggggttt 60
173 tgccactgcc agagggacgt ctcagacttc atcttcccaa atcttggcag atcaccccct 120
175 tagcagggtc tgcacatctc agccggg atg aag ctg gtt ccc gta gcc ctc atg 174
176                             Met Lys Leu Val Pro Val Ala Leu Met
177                                                 -90
179 tac ctg ggc tcg ctc gcc ttc ctg ggc gct gac aca gct cgg ctc gac  222
180 Tyr Leu Gly Ser Leu Ala Phe Leu Gly Ala Asp Thr Ala Arg Leu Asp
181 -85                 -80                 -75                 -70
183 gtg gcg gca gag ttc cga aag aaa tgg aat aag tgg gct cta agt cgt  270
184 Val Ala Ala Glu Phe Arg Lys Lys Trp Asn Lys Trp Ala Leu Ser Arg
185                 -65                 -60                 -55
187 gga aaa aga gaa ctt cgg ctg tcc agc agc tac ccc acc ggg atc gcc  318
188 Gly Lys Arg Glu Leu Arg Leu Ser Ser Ser Tyr Pro Thr Gly Ile Ala
189                 -50                 -45                 -40
191 gac ttg aag gcc ggg cct gcc cag act gtc att cgg ccc cag gat gtg  366
192 Asp Leu Lys Ala Gly Pro Ala Gln Thr Val Ile Arg Pro Gln Asp Val
193             -35                 -30                 -25
195 aag ggc tcc tct cgc agc ccc cag gcc agc att ccg gat gca gcc cgc  414
196 Lys Gly Ser Ser Arg Ser Pro Gln Ala Ser Ile Pro Asp Ala Ala Arg
197    -20                 -15                 -10
199 atc cga gtc aag cgc tac cgc cag agt atg aac aac ttc cag ggc ctg  462
200 Ile Arg Val Lys Arg Tyr Arg Gln Ser Met Asn Asn Phe Gln Gly Leu
201 -5                  -1  1           5                   10
203 cgg agc ttc ggc tgt cgc ttt ggg acg tgc acc gtg cag aag ctg gcg  510
204 Arg Ser Phe Gly Cys Arg Phe Gly Thr Cys Thr Val Gln Lys Leu Ala
205             15                  20                  25
207 cac cag atc tac cag ttc acg gac aaa gac aag gac ggc gtc gcc ccc  558
```

```
208 His Gln Ile Tyr Gln Phe Thr Asp Lys Asp Lys Asp Gly Val Ala Pro
209         30              35              40
211 cgg agc aag atc agc ccc cag ggc tac ggc cgc cgg cgc cga cgc tct  606
212 Arg Ser Lys Ile Ser Pro Gln Gly Tyr Gly Arg Arg Arg Arg Arg Ser
213         45  .             50              55
215 ctg ccc gaa gcc agc ctg ggc cgg act ctg agg tcc cag gag cca cag  654
216 Leu Pro Glu Ala Ser Leu Gly Arg Thr Leu Arg Ser Gln Glu Pro Gln
217  60              65              70              75
219 gcg cac ggg gcc ccg gcc tcc ccg gcg cat caa gtg ctc gcc act ctc  702
220 Ala His Gly Ala Pro Ala Ser Pro Ala His Gln Val Leu Ala Thr Leu
221             80              85              90
223 ttt agg att taggcgccta ctgtggcagc agcgaacagt cgcgcatgca          751
224 Phe Arg Ile
226 tcatgccggc gcttcctggg gcgggggct tccggagcc gagccctca gcggctgggg 811
228 cccgggcaga gacagcattg agagaccgag agtccgggag gcacagacca gcggcgagcc 871
230 ctgcatttc aggaacccgt cctgcttgga ggcagtgttc tcttcggctt aatccagccc 931
232 gggtccccgg gtggggggtgg agggtgcaga ggaatccaaa ggagtgtcat ctgccaggct 991
234 cacggagagg agaaactgcg aagtaaatgc ttagacccc agggggcaagg gtctgagcca 1051
236 ctgccgtgcc gcccacaaac tgatttctga aggggaataa ccccaacagg gcgcaagcct 1111
238 cactattact tgaactttcc aaaacctaga gaggaaaagt gcaatgtatg ttgtatataa 1171
240 agaggtaact atcaatattt aagtttgttg ctgtcaagat ttttttttgt aacttcaaat 1231
242 atagagatat ttttgtacgt tatatattgt attaagggca ttttaaaaca attgtattgt 1291
244 tccccctcccc tctattttaa tatgtgaatg tctcagcgag gtgtaacatt gtttgctgcg 1351
246 cgaaatgtga gagtgtgtgt gtgtgtgtgc gtgaaagaga gtctggatgc ctcttgggga 1411
248 agaagaaac accatatctg tataatctat ttacataaaa tgggtgatat gcgaagtagc 1471
250 aaaccaataa actgtctcaa tg                                          1493
253 <210> SEQ ID NO: 4
254 <211> LENGTH: 188
255 <212> TYPE: PRT
256 <213> ORGANISM: Sus scrofa
258 <400> SEQUENCE: 4
259 Met Lys Leu Val Pro Val Ala Leu Met Tyr Leu Gly Ser Leu Ala Phe
260             -90             -85             -80
262 Leu Gly Ala Asp Thr Ala Arg Leu Asp Val Ala Ala Glu Phe Arg Lys
263             -75             -70             -65
265 Lys Trp Asn Lys Trp Ala Leu Ser Arg Gly Lys Arg Glu Leu Arg Leu
266         -60             -55             -50
268 Ser Ser Ser Tyr Pro Thr Gly Ile Ala Asp Leu Lys Ala Gly Pro Ala
269     -45             -40             -35
271 Gln Thr Val Ile Arg Pro Gln Asp Val Lys Gly Ser Ser Arg Ser Pro
272 -30             -25             -20             -15
274 Gln Ala Ser Ile Pro Asp Ala Ala Arg Ile Arg Val Lys Arg Tyr Arg
275             -10             -5              -1  1
277 Gln Ser Met Asn Asn Phe Gln Gly Leu Arg Ser Phe Gly Cys Arg Phe
278         5               10              15
280 Gly Thr Cys Thr Val Gln Lys Leu Ala His Gln Ile Tyr Gln Phe Thr
281     20              25              30
283 Asp Lys Asp Lys Asp Gly Val Ala Pro Arg Ser Lys Ile Ser Pro Gln
284 35              40              45              50
```

```
                    RAW SEQUENCE LISTING              DATE: 02/08/2002
                    PATENT APPLICATION: US/10/030,298 TIME: 10:52:14

                    Input Set : A:\SO043PCTSEQ.txt
                    Output Set: N:\CRF3\02082002\J030298.raw

    286 Gly Tyr Gly Arg Arg Arg Arg Arg Ser Leu Pro Glu Ala Ser Leu Gly
    287               55              60                  65
    289 Arg Thr Leu Arg Ser Gln Glu Pro Gln Ala His Gly Ala Pro Ala Ser
    290           70              75                  80
    292 Pro Ala His Gln Val Leu Ala Thr Leu Phe Arg Ile
    293           85              90
    296 <210> SEQ ID NO: 5
    297 <211> LENGTH: 1376
    298 <212> TYPE: DNA
    299 <213> ORGANISM: Rattus norvegicus
    301 <220> FEATURE:
    302 <221> NAME/KEY: CDS
    303 <222> LOCATION: (154)..(708)
    305 <220> FEATURE:
W--> 306 <221> NAME/KEY: mat peptide
    307 <222> LOCATION: (433)..(582)
    309 <400> SEQUENCE: 5
    310 tccagccttt accgctcctg gtttctcggc ttctcatcgc agtcagtctt ggactttgcg 60
    312 ggttttgccg ctgtcagaag gacgtctcgg actttctgct tcaagtgctt gacaactcac 120
    314 cctttcagca gggtatcgga gcatcgctac aga atg aag ctg gtt tcc atc gcc 174
    315                                     Met Lys Leu Val Ser Ile Ala
    316                                                     -90
    318 ctg atg tta ttg ggt tcg ctc gcc gtt ctc ggc gcg gac acc gca cgg 222
    319 Leu Met Leu Leu Gly Ser Leu Ala Val Leu Gly Ala Asp Thr Ala Arg
    320     -85             -80                 -75
    322 ctc gac act tcc tcg cag ttc cga aag aag tgg aat aag tgg gcg cta 270
    323 Leu Asp Thr Ser Ser Gln Phe Arg Lys Lys Trp Asn Lys Trp Ala Leu
    324 -70                 -65                 -60                 -55
    326 agt cgt ggg aag agg gaa cta caa gcg tcc agc agc tac cct acg ggg 318
    327 Ser Arg Gly Lys Arg Glu Leu Gln Ala Ser Ser Ser Tyr Pro Thr Gly
    328                 -50                 -45                 -40
    330 ctc gtt gat gag aag aca gtc ccg acc cag act ctt ggg ctc cag gac 366
    331 Leu Val Asp Glu Lys Thr Val Pro Thr Gln Thr Leu Gly Leu Gln Asp
    332             -35                 -30                 -25
    334 aag cag agc acg tct agc acc cca caa gcc agc act cag agc aca gcc 414
    335 Lys Gln Ser Thr Ser Ser Thr Pro Gln Ala Ser Thr Gln Ser Thr Ala
    336         -20                 -15                 -10
    338 cac att cga gtc aaa cgc tac cgc cag agc atg aac cag ggg tcc cgc 462
    339 His Ile Arg Val Lys Arg Tyr Arg Gln Ser Met Asn Gln Gly Ser Arg
    340     -5              -1  1               5                   10
    342 agc act gga tgc cgc ttt ggg acc tgc aca atg cag aaa ctg gct cac 510
    343 Ser Thr Gly Cys Arg Phe Gly Thr Cys Thr Met Gln Lys Leu Ala His
    344             15                  20                  25
    346 cag atc tac cag ttt aca gac aaa gac aag gac ggc atg gcc ccc aga 558
    347 Gln Ile Tyr Gln Phe Thr Asp Lys Asp Lys Asp Gly Met Ala Pro Arg
    348         30                  35                  40
    350 aac aag atc agc cct caa ggc tat ggc cgc cgg cgc cgg cgt tcc ctg 606
    351 Asn Lys Ile Ser Pro Gln Gly Tyr Gly Arg Arg Arg Arg Arg Ser Leu
    352         45                  50                  55
```

VERIFICATION SUMMARY   DATE: 02/08/2002
PATENT APPLICATION: US/10/030,298   TIME: 10:52:15

Input Set : **A:\SO043PCTSEQ.txt**
Output Set: **N:\CRF3\02082002\J030298.raw**

L:9 M:270 C: Current Application Number differs, Replaced Application Number
L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:30 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:1
L:167 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:3
L:306 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:5